DR. SKINNER: Could I comment? I believe there are 20 deaths, and that was one of the questions I was concerned about because an average age of 53 years, 20 deaths in five years in that age group sounds like a lot a deaths.

PANEL CHAIRPERSON NAIDU: Thank you, Dr. Skinner.

MS. MARLOW: As far as the number of -- I'm sorry. My name is Marie Marlow. I'm a consultant to Smith, and Nephew and as such I'm paid for work I do for them.

In your panel packets, there are narratives for each of the deaths that occurred in this trial, the reasons for them, length of time after surgery. We can get a summary table together for you and discuss that in greater detail if you'd like.

DR. BLUMENSTEIN: How were the deaths handled with respect to the estimation of survivorship?

MS. MARLOW: I'll pass that to a statistician.

MR. DeMUTH: Yes. Actually -- George

DeMuth -- we censored at the time of death if there was a death available. So if there was a revision, well, if they had revised the packet in the revision, but we just censored at the time of death those patients.

PANEL CHAIRPERSON NAIDU: Thank you.

DR. SKINNER: Could I follow up on that, please, Dr. Naidu?

PANEL CHAIRPERSON NAIDU: Sure.

DR. SKINNER: I'd be interested to know what the number of expected deaths in a group of 2,385 individuals would be over the following five years with an average age of 53 to see if that's actually an expected death number of 20, which is the number of deaths in it.

I looked at the packet. It doesn't look like any of them were caused by the hip. so it seems like an awful lot of deaths in a group of 53 year old individuals.

DR. RICHARDSON: My name is James Richardson, Professor of Orthopedics at Oswestry, U.K.

I'm an orthopedic surgeon. I've been

studying hip resurfacing for eight years now, and I'm not affiliated with Smith & Nephew. I have got funding in order to carry out my research, but I have no personal benefit.

I've been interested in this question, and using the data from the Outcome Center to try and find a solution to this. The average death rate of 20 out of 2,385, you can check my mathematics, but it's in the order of, as a percentage, one percent, but that's over three years, the average 3.3 years. Again, I'm just giving general figures and I will explain why.

The death rate per year then comes to about .3 percent. Now, studying death rates is tricky because each age band, I'm afraid we all have a different expectation of death in each year, but for the age group within the U.K., 55 to 59, information I have is that for ladies who are the tougher of the species, the expected death rate is .47 percent, whereas for us feeble men it's .74, almost twice.

So per year, you can see that the death rates in the hip replacement group are actually slightly below what might be predicted.

I qualify my comments by saying that if you really want to analyze this in detail, then looking at simple averages it not sufficient. You really need to break down all the data and look at the death rates for each age band and it becomes more complicated.

But at least from what I can find out, I don't think it's a major concern.

Thank you.

PANEL CHAIRPERSON NAIDU: Thank you.

Dr. Mayor, you had a question?

DR. MAYOR: Yes, I have a couple of questions for Derek McMinn.

Congratulations for a prodigious body of work, but I wonder if you could clarify a couple of issues for me. You suggested that revision is a fairly straightforward undertaking, although it might not be the most enthusiastic problem that the patient might perceive. How do you handle the acetabulum in the process of doing the revisions that you've done, particularly, say, for femoral neck fracture?

And then the second question, if I can

present them both together, I was confused by the description of femoral head collapse and identified as distinction from avascular necrosis and wondered how that distinction was drawn.

DR. McMINN: I'll try to bring up the slide.

MS. MARLOW: Chairman Naidu, with your permission, may I ask Dr. Mayor a question?

PANEL CHAIRPERSON NAIDU: Yes.

MS. MARLOW: Thank you.

Dr. Mayor, in addition to Mr. McMinn, Dr. Rorabeck is here because he has experiences with these cases in Canada also. So if you would like another perspective in addition to Mr. McMinn's, please feel free to ask Dr. Rorabeck about his experience.

DR. MAYOR: I am always happy to have Dr. Rorabeck's perspective.

DR. McMINN: Derek McMinn again.

On the question of what we do with the acetabulum in the event of, let's say, a femoral neck fracture, the acetabulum is left alone, provided the acetabulum is in a good position and at surgery

there's no adverse effect like, for example, evidence of age loading or wear of the component.

So you put in cemented or cementless stem, according to your preference as a surgeon, and then put a modular ball on it that matches the existing cup. So that's quite straightforward.

I'm happy to answer anything else on that if you have further questions.

Trying to differentiate between a collapsed head and avascular necrosis, that's a good question because it's a bit like trying to identify the cause of a fire when you house has burned down. I've been there 20 years ago, and when you're going through the ashes, it's quite difficult and tough to decide on the cause of the fire.

If you get collapse of the head, then it's the devil's own job to differentiate between collapse from, let's say, squashing of a lot of cysts or avascular necrosis that has occurred.

The only ones where you're pretty clear that it's avascular necrosis is if no complete collapse has occurred, and histology then shows you a

segment of the head distal to the collapsed area that has avascular necrosis changes in it.

But I absolutely agree that some of these cases could have been categorized in avascular necrosis or collapsed head, and when the collapse has occurred, it's very difficult to know what the original pathology was.

DR. MAYOR: Thank you.

PANEL CHAIRPERSON NAIDU: Thank you, Dr. McMinn.

Any other questions from the panel? Dr. Skinner.

DR. SKINNER: Harry Skinner.

I wanted to also ask Mr. McMinn some questions if he could answer.

First of all, could I follow up on that femoral neck fracture thing?

I went through the data, and it looks like most of the femoral neck fracture, head collapse things happened earlier on since '97, and the package insert is going to address that by limiting the indications or contraindications to less than 50

percent of femoral head involvement in avascular necrosis.

Do you think or does the data show that the number of femoral head problems has decreased since you have sort of instituted those changes in your selection criteria?

DR. McMINN: It's true -- Derek McMinn again -- it's true what you say about avascular necrosis being a major problem for femoral head collapse. The highest percentage group of all in our femoral head collapses are avascular necrosis. So in those with a preexisting diagnosis of avascular necrosis, they had a four percent collapse rate.

So if you have extensive AVN and want to do a resurfacing, what we have subsequently found out is that whether you get collapse or not (a) depends on the magnitude of the original femoral head lesion and (b) whether the pathology is recurring or not.

So if you had, for example, a traumatic dislocation and it was a once an event time, then the chances of a subsequent collapse of that head are small.

However, if you have avascular necrosis caused by, for example, alcoholics, increased alcohol intake, they have a bad post op record with further collapse of their heads. Mind you they also have a bad postop record with total hip replacement and falling out of bed and dislocating their hips, et cetera.

There was one other point which I've just forgotten. Could you remind me of the other point in your question?

DR. SKINNER: I think you've already answered it. You think that the changes in the labeling basically, the indications/contraindications, will address this because you've changed your criteria for selection, I gathered from the data given to us.

So you think that the incidence of femoral neck fracture collapse is going to be decreased as time goes on, at least in your series.

DR. McMINN: The answer to that is it has reduced with the passage of time, but that really is based on the light of experience and understanding that there's no point in trying to attempt the

impossible, and if you try and fix onto a femoral head remnant with virtually no bone in it, it's going to fall to bits.

So I doubt a break with a sense of realism as I grew older took part here.

DR. SKINNER: Another question. I would like to get a sense of your practice of orthopedic surgery to get an idea of what the selection process for these 1,626 patients was, 2,385 patients. Do the patients come into your place, you select them. They either have enough arthritis to have surgery and you do a surface replacement on them or if they don't have enough and you send them away or they have enough and you decide that they're a good candidate for a surface replacement or they're a bad candidate, and then you do a total hip on them.

I mean, how does that go?

And a second thing. Is your practice a referral practice with only non -- I'm not sure about the British system -- but non-health service patients or whatever the practice is? Is there a selection process there?

DR. McMINN: Okay. Mine is mainly a referral practice, but I do get through family practitioners referred local patients. So typically in my clinic out of 20 patients I may see, 18 of them would be referrals from outside my city, typically referred by another orthopedic surgeon because the patient was young, active, had good bone stock, and wanted a resurfacing.

So a lot of the patients that I see have come specifically for a Birmingham hip resurfacing sent there by their orthopedic surgeon, and quite a number come from other countries. For example, we were looking through the data. Sixty-five of my patients included in the information are from North America. So it's an international practice.

However, if I get the local family doctor sending a patient from Edgbaston, particularly a lady in her 70s with an arthritic hip, the question of hip resurfacing will never get mentioned, and I will do a total hip replacement on her without discussing what my main activity is, namely, hip resurfacing. Because that's on a particular group of patients who are

younger, more active, and are likely to cause a failure of a total hip replacement.

DR. SKINNER: Roughly how many total hips have you done in the last seven years?

DR. McMINN: I would do somewhere between 50 and 100 hips a year. That number has decreased as I've got a little older and not doing quite as much, and my younger colleagues, I tend to try and pass on total hips, and particularly revision total hips to, but resurfacing, I still enjoy doing that. So I don't pass those on.

DR. SKINNER: Now, a little bit of the clinical data. I noticed in your data that you had a fairly high pain rate. Something like 15 percent of your patients reported pain compared to some of the ceramic studies. There was quite a bit lower rate of pain in those.

Is there a particular reason for that?

DR. McMinn: Well, you have to understand that this is a review of my notes. You're holding the microphone. Do you want to interrupt me?

(Laughter.)

DR. McMinn: This is a review of my notes by an outside group of consultants who I understand were detailed to go through that with a fine toothed comb and record everything.

Now, when you look longitudinally at the data, it's pretty bizarre because postoperatively small percentage of patients recorded pain. Was there any mention of pain? Yet at a year, there was a higher percentage, and you'd think that's completely bizarre.

But this was from the records and there was only mention made of pain if the pain was worse than either the doctor or the patient anticipated. So that's why postoperatively there's a low figure, percentage, for pain and later on there's a higher percentage for pain.

In other words, the patients said, "Hey, what's this pain?" So any level of pain that was recorded.

DR. SKINNER: Well, I happen to think that the 15 percent is more like a realistic number, to be honest, than the three percent or whatever, but how

about limp? A fair number of your patients also seem to have a limp even I think a year out it was.

MS. MARLOW: Again, Marie Marlow.

We were the company that Smith & Nephew -- sorry. I'm involved with the company that did the audit for Smith & Nephew of these data, and when we deployed the auditors, we told them that since this is a retrospective review, that every single incident had to be recorded.

What we did also was combined comments that we found in the Oswestry database, along with the comments that we found in Mr. McMinn's series. So there's an area on the Oswestry form, for example, for a patient to make notes, offer comments. If a patient made a comment in that field, we didn't censor it; we didn't filter it. If they said, "I have a mild limp," we recorded that as an adverse event.

DR. SKINNER: Okay.

MS. MARLOW: Is that helpful?

DR. SKINNER: Yes, that's very helpful. Thank you.

PANEL CHAIRPERSON NAIDU: Thank you, Dr.

Skinner. Thank you, Dr. McMinn.

Yes, Ms. Whittington.

MS. WHITTINGTON: I have a question. There seemed to be an extremely high number of patients with wound exudate, and if that data is taken from office notes after the fact, that's even more disconcerting to me since that's a primary indicator for potential late infection.

MS. MARLOW: Again, Marie Marlow.

Thanks for that question. Part of the office notes or part of the records in the patient files at the McMinn Center included the discharge notes, as well as the op notes. So if there is a comment on the discharge note about a patient being instructed as to dressing changed, wound care, that was listed as an adverse event because, again, the auditors were instructed not to filter anything, not to make interpretations about whether something was an adverse event or not. If there was a comment there, it would have recorded it.

MS. WHITTINGTON: Was there a correlation to those patients with the patients who had revision

or later infections as noted in the --

MS. MARLOW: I'm sorry. I didn't answer your question about the late one.

MS. WHITTINGTON: Right.

MS. MARLOW: Postop the cutoff was 30 days. So the column that you're looking at in the adverse events table, that's a 30-day cutoff. Anything after 30 days got moved into the one-year column.

MS. WHITTINGTON: Okay. Was there a correlation to those patients to the patients who did develop late wound infections and had some of them revisions that looked like later?

MS. MARLOW: I don't know from the data I have in front of me. I'll see if we can get that answer for you.

MS. WHITTINGTON: Okay.

PANEL CHAIRPERSON NAIDU: Thank you.

Why don't we take a short break now for 15 minutes and we'll come back with the FDA presentation then?

Thank you.

(Whereupon, the foregoing matter went off the record at 10:46 a.m. and went back on the record at 11:03 a.m.)

PANEL CHAIRPERSON NAIDU: We will now have the FDA presentation on this PMA. The first FDA presenter is Mr. John Goode, the review team leader for this PMA. He will introduce the other FDA presenter.

Mr. Goode.

MR. GOODE: Thank you, Dr. Naidu.

Good morning. My name is John Goode. I'm a biomedical engineer and reviewer in the Orthopedic Devices Branch and the lead reviewer for the Smith & Nephew premarket approval application for the Birmingham hip resurfacing system. I will be presenting the device description, preclinical and clinical information, and FDA statistician Dr. Chang Lao will summarize the statistical information in the PMA.

But first I would like to identify some of the reasons for the panel meeting and the topics for which we are seeking panel input.

The Birmingham hip resurfacing system, or BHR, is a first of a kind device in the United States. That is the first total hip system with a resurfacing femoral component and metal-on-metal articulating surfaces.

The PMA is supported by clinical data essentially from one source, the surgical experience of Dr. Derek McMinn, who implanted devices primarily at the Birmingham Nuffield Hospital, City of Birmingham, United Kingdom.

The PMA includes safety and effectiveness data from an uncontrolled case series of all 2,385 procedures implanted with the BHR device from July 1997 through May 2004.

FDA requests expert clinical opinion on the following topics: the way in which the safety and effectiveness data were collected; the results of the study; and the applicability of data collected outside of the United States by a single investigator to the target U.S. population, U.S. practice of medicine, and U.S. orthopedic surgeon population.

This slide includes the outline for the

rest of FDA's presentation. It is our goal to summarize the information in the PMA to help the panel address FDA's questions. I will briefly present the device description and preclinical testing information. I will then discuss the sponsor's clinical data, focusing on the way in which the patients were selected to receive the BHR, the indications for use, the way in which the data was collected.

Then I will present a proposed post approval study. Dr. Chang Lao will then summarize the statistical information in the PMA and that will conclude FDA's presentation.

After lunch, I will present seven FDA questions for panel discussion.

First, the device description. I believe that the sponsor has adequately summarized the device description in their presentation, and I just had one clarification regarding the acetabular shells. There are three styles of acetabular shell. One is the standard cup. The other is the dysplasia cup, and a third being the bridging cup, and the dysplasia and

bridging cups are both for dysplasia indications with the bridging cup just being slightly thicker than the dysplasia cup.

I believe the rest of the information was adequately covered by the sponsor. The sponsor provided preclinical information that included the evaluations listed on this slide and others. Some of this information was summarized by the sponsor in their presentation, and FDA believes the sponsor has adequately addressed the preclinical issues for the BHR device.

Now I'll discuss the way in which the patients were selected for this study. Please note that one of the FDA questions will ask for your comments on the sponsor's device labeling.

Generally, prospective clinical investigations predefined the study population with specific inclusion and exclusion criteria. This theoretically allows the study results to be generalized to that diagnostic group. In case series studies it is more difficult to generalize the results to a defined population because the patients were not

enrolled for predefined conditions.

This is the case for the clinical data provided in this PMA submission. The clinical data were derived primarily from the surgical experience of a single surgeon. This surgeon did not predefine a set of diagnostic indications for the device, but instead provided a list of the diagnostic indications for the patients implanted with the device.

During the same time period Dr. McMinn implanted the BHR devices, he also had patients who either had no surgery or conventional total hip replacement. However, a complete review of these patients was not presented in the PMA. With this information, it might have been possible to retrospectively determine what criteria, if any, were used to select candidates for the BHR.

As an alternative and in order to retrospectively develop the indications for use in physician labeling, the sponsor provided a list of the factors that contributed to Dr. McMinn's decision to perform a total hip replacement, or a THR, in certain patients rather than the BHR hip resurfacing

procedure.

These factors included advanced age. Patients of an advanced age, especially those with low activity levels, were typically candidates for THR rather than BHR. Only 8.1 percent of the 2,385 cases were greater than 65 years of age. In these cases the BHR were selected despite advanced age if the patients had high activity levels and had good bone stock of the femoral head.

Low activity level. Patients with a low activity level were considered at lowered risk for future revision and, therefore, good candidates for THR and not BHR. Low activity level was characterized by no participation in sports activities, no heavy work required by job, a sedentary or retired lifestyle, or co-morbidities that precluded a high activity level, such as severe arthritis in other joints or severe VFR disease.

Poor bone stock. Patients with poor bone stock were selected for THR rather than BHR because they were considered at risk for femoral neck fracture or femoral head collapse. With a resurfacing

procedure poor bone stock was retrospectively defined and is a contraindication for the BHR.

The sponsor stated that Dr. McMinn's preoperative evaluation was typically sufficient to screen candidates for BHR versus THR, and that only in rare instances would the planned surgical procedure be revised intraoperatively. Because of the potential for a change in the preop plan, patients were consented for both a BHR and THR procedure.

Based upon the population studied and the factors just mentioned and also an analysis of the BHR revisions, which included femoral neck fracture, femoral head collapse, dislocation, AVN, and infection, the sponsor proposed the following indications for use for the device.

The BHR system is intended for patients requiring primary hip resurfacing due to noninflammatory arthritis and inflammatory arthritis, such as rheumatoid arthritis. The BHR hip resurfacing arthroplasty is intended for joint replacement in patients who are at risk of requiring future ipsilateral hip joint revision.

While it's impossible to predict if a patient will require more than one joint replacement, several factors are known to increase risk of revision surgery, including age less than 55 years at index surgery and/or high physical activity level postoperative.

These are the contraindications for the device, and I believe the sponsor has covered these, focus particularly on one bullet item, which was regarding the inadequate bone stock, which includes severe osteopenia, osteonecrosis or avascular necrosis, with greater than 50 percent involvement of the femoral head and multiple cysts on the femoral head greater than one centimeter.

Also, females of childbearing age due to unknown effect of a fetus on the metal ion release.

In addition to the factors described above, the sponsor also considered a review of 50 BHR femoral neck fractures reported by Schimmin and Back in the development of the labeling. In this publication, the authors reported a review of 3,497 BHR cases performed in Australia by 89 surgeons.

There were 50 femoral neck fractures in the series, or 1.46 percent, which the authors attribute to osteoporosis and difficulties in implantation of the head and the cup leading to notching of the superior femoral neck, varus placement of the device by more than five degrees, difficulty in interoperative alignment, impaction of the femoral component, and poor exposure.

Based upon these findings, the sponsor added the following warnings and precautions to the labeling. "Warning: avoid notching the femoral neck as this may lead to femoral neck fracture. Avoid placing the femoral component in varus. Varus placement of the femoral component has been associated with femoral neck fracture," and the following precaution: "Improper selection, placement, positioning, and fixation of the implant component may result in early implant failure."

The objective of this PMA is to demonstrate the safety and effectiveness of the BHR system. The safety assessments included data on revisions, adverse events, and a metal ion literature

review.

Effectiveness assessments included survivorship, radiographic data, pain and function data as evaluated by the Oswestry modified Harris Hip Score and patient satisfaction data.

Before I go into the description of the results, I quickly wanted to describe for you and paraphrase what valid scientific evidence is. This information was provided to the panel members as a part of their training, and I just wanted to paraphrase some particular parts of what the FDA considers to be valid scientific evidence. This is just being added to my talk at this particular time.

Again, you can read the entire statement of what valid scientific evidence is in your packet, but I just wanted to summarize these parts quickly before I present the clinical data.

Valid scientific evidence is evidence from well controlled investigations, partially controlled studies, studies and objective trials without matched controls, well documented case histories conducted by qualified experts, and reports of significant human

experience with a marketed device. That is what valid and scientific evidence is.

Valid and scientific evidence is not isolated case reports, random experience, reports lacking sufficient details to permit scientific evaluation, and unsubstantiated opinions are not recorded as valid scientific evidence.

Again, you have the entire definition in your packet, and it is a question for the panel today on whether or not you believe this to be valid scientific evidence.

I will now go into a description of the study. As the sponsor described, there were 2,385 BHR procedures, and they were divided into the following three main cohorts. Again, an X-ray cohort which is the first 124 BHR cases implanted in 1997; the Oswestry cohort, which was the next 1,502 cases; and the McMinn cohort, which was the next 759 cases.

Note that there were five cases in the McMinn cohort whose implantations were performed prior to April 2002. These cases should have been part of the Oswestry cohort, but for unknown reasons were not.

Therefore, unlike the majority of the McMinn cohort, some of these five cases have longer term follow-up.

Where there were common data elements collected in these three cohorts, the sponsor pooled this information into two combined cohorts, which included what we're calling the overall McMinn cohort or the combination of the X-ray, the Oswestry, and the McMinn cohorts, as well as the X-ray-Oswestry combined cohort.

The overall McMinn cohort contributed to the assessment of safety, including adverse events and revisions. The X-ray cohort contributed to the assessment of radiographic effectiveness. Radiographic evaluations were not provided for the 1,502 procedures in the Oswestry cohort or the 759 procedures in the McMinn cohort.

The X-ray and Oswestry combined cohort contributed to the assessment of survivorship and patient satisfaction, and the 1,111 unilateral procedures in this combined cohort contributed to the assessment of pain and function effectiveness data as evaluated by the Oswestry modified Harris Hip or OSHIP

Score.

Note that the pain and function data for the 759 procedures in the McMinn cohort were collected using the Oxford Hip Score evaluation method and not the OSHIP score. The sponsor explained that because these data were not tracked by the Oswestry Outcomes Center, but by the National Health Services Center, the sponsor did not have access to the Oxford Hip Score data.

The main data sources were just presented, but the sponsor also included additional, less complete data on 3,374 BHR cases performed by 140 surgeons worldwide other than Dr. McMinn. The follow-up for these cases also was contracted with the Oswestry Outcomes Center and includes primarily the same data as that provided for the X-ray and Oswestry combined cohort. The Oswestry Outcomes Center has provided Smith & Nephew access to all available data for the BHR cases from its database.

Although the sponsor considers the data from this additional cohort to be of some value, Smith & Nephew has no ability to independently verify any of

the data provided to the Oswestry Outcomes Center by sites other than the McMinn Center and Dr. McMinn and has no ability to request additional follow-up or clarifications of any kind from non-McMinn patients or physicians.

For these reasons, this data has some limitations and is not considered a primary data source for this PMA.

Now I will summarize the way in which the safety data was collected in this study. Please note that one of the FDA questions will ask for your comments on the reliability of these data collection methods.

The safety data, including adverse events and revisions, were collected by the following three methods: the Oswestry Outcomes Center using an annual patient completed mail-in questionnaire, the McMinn Center by recording the findings of postoperative patient visits, and recording information provided to Dr. McMinn by primary care physicians.

Dr. McMinn's follow-up was described as follows. Dr. McMinn performed regular evaluations

which included history, physical examination, radiographs to assess implant status, and any necessary laboratory work in the preoperative and postoperative time periods according to standard practice, although the time points and evaluations were not according to standard protocol.

All revision surgeries were performed by Dr. McMinn, except in one known case. Therefore the revision status was directly known to Dr. McMinn.

There were no predefined follow-up time windows, standardized clinical evaluations, adverse event report forms or standardized radiographic evaluations.

The sponsor also provided the following information regarding the follow-up procedure or what I'm going to call the OOC. The OOC collected safety data on revisions and adverse events, again using an annual patient completed mail-in questionnaire. With the exception of eight cases who withdrew or did not agree to participate in this study, all other cases are not considered lost to follow-up since the OOC continues to make attempts to contact patients.

Of the 180 cases missing, the last theoretical expected mail-in questionnaire follow-up, 84 are missing on at least two-year reevaluations, while 96 are only missing their last evaluation. These 180 cases represent 11 percent of the Oswestry/X-ray combined cohort.

The OOC identified several steps taken to regain contact if a patient does not respond to a request for information, including sending reminder letters, E-mail or phoning the patient, contacting the consulting surgeon by letter or telephone, using a national strategic tracing service database and Internet census information to determine the patient's whereabouts.

If the patient is still not found, an additional request for information is sent to the patient's last known address via Royal Mail and E-mail until the tenth anniversary of the operation.

Patients are not classified as lost to follow-up until all avenues have been exhausted. The sponsor stated that they performed a 100 percent audit of all 2,385 procedures in the overall McMinn cohort

and, therefore, believe that all reported adverse event information has been captured.

In addition to the safety date collection methods outlined above, the sponsor provided a metal ion literature analysis. Included in the sponsor's analysis was an unpublished report by Daniel Zee and McMinn. The authors conducted four metal ion studies, and I believe the sponsor has adequately summarized those studies in their presentation.

Now I will summarize the way in which the effectiveness data was collected. Please note that one of the FDA questions will ask for your comments on the reliability of these data collection methods.

The primary effectiveness measurement was survivorship for procedures with a minimum of two years postop. Of these 1,626 procedures, data were available on 546 of the 601 BHR procedures eligible for five-year follow-up or 90.8 percent.

The data for the survivorship study was collected using the same methods presented for the safety data that included the OOC and the McMinn Centers. The PMA also contained the results of an

independent radiographic review of the X-ray cohort, which was the first 124 procedures performed. The sponsor adequately summarized the radiographic information and, again, comparison to baseline films were made for each of the 108 cases that were out to five years in the postoperative time period, and the baseline films were, again, usually within three months, but eight of those 108 procedures the baseline films were evaluated between 110 and 860 days postoperative.

The radiographs were independently evaluated by Dr. Nick Evans. A prospective protocol was used to assess the radiographs. The five-year AP and lateral view radiographs were compared to baseline radiographs for migration, acetabular orientation, radiolucency, heterotopic ossification, and other radiographic findings.

A radiographic success was defined as having all of the following: absence of radiolucencies or radiolucency in any one or two zones, component migration less than or equal to two millimeters, and a change in acetabular angle less

than five degrees.

A radiographic failure was defined as the presence of incomplete or complete radiolucencies or radiolucency in all zones and migration of the component greater than two millimeters or a change in the acetabular orientation less than or equal to five degrees.

Pain function and movement data were collected by the OOC using an annual patient completed, mail-in questionnaire. The patient responses to the questions were used to generate the Oswestry modified Harris Hip or OSHIP score. I believe that this is the first time the FDA has evaluated the OSHIP scoring system in a marketing application. Therefore, the FDA has asked the sponsor on how the OSHA data were collected, how the OSHIP scoring system was developed, and asked for a justification for its use.

The OSHIP questionnaire allows patient assessments without direct physician evaluation. No other sources of pain and function information were used to support this PMA.

The sponsor summarized the OOC standard operating procedure for data input and clarification of the patient administered OSHIP questionnaires. Any questionnaire with missing, unclear, or conflicting information was returned to the patient with specific instructions for completing the form. The preferred method was by mail. However, E-mail and telephone were also used to complete these questionnaires.

If the data were not collected, the score from any missing item was assumed to be the lowest possible, which was typically zero.

Now I'll discuss how the OSHIP patient questionnaire was developed. It was developed by Professor James Richardson with the following premises. He believed that a long-term evaluation following hip replacement is essential. Follow-up must be regular and large samples are necessary. Long-term and large sample follow-up is difficult to obtain when using a score that requires a surgeon or radiologist assessment.

Physician administered surveys are susceptible to bias which may inflate the final score

and may not truly represent the patient's own feelings, and questionnaires needed to be simple and relatively short to make long-term and large scale collection of data more efficient.

Building on these premises, Professor Richardson developed the OSHIP scoring system by combining elements of both the Harris Hip Score and Merle d'Aubigne scores. As presented, the OSHIP produces an overall index score similar to that of the Harris Hip Score between zero and 100. The OSHIP score is made up of three domains of pain, function, and hip movement. The main difference between the OSHIP questionnaire and the Harris Hip Score is that the OSHIP allows patient assessments without direct physician evaluation.

In addition, the OSHIP questionnaire does not include the three HHS questions regarding a physician assessment of range of motion, which is five points, absence of deformity, which is four points, and the patient's ability to put on shoes and put on socks and tie shoes, which is four points, but substitutes a movement question which is 13 points

that is intended for the patient to estimate their ability to flex their hip.

There are additional differences between the OSHIP and Harris Hip Scores in the phrasing of some of the questions and the point values that correspond to some of the answers.

Again, FDA requested that the sponsor justify the use of the OSHIP scoring system, and also the validity of patient self-administered questionnaires, and the sponsor summarized several literature references which I'm now going to discuss in some detail.

While a paper by Ragab and co-workers reported a lack of correlation between patient self-assessment of pain and function and physician assessment of pain and function with a correlation of .467, several other researchers have reported the opposite, a very close correlation between patient self-assessment and physician assessment.

Research by Mohammed and co-workers demonstrated that patients are able to accurately respond to Harris Hip Score questions regarding pain

and function with little difficulty, and there is excellent correlation between overall HHS pain and function scores reported by the patient and the overall HHS pain and function scores reported by physicians with a correlation of .99.

In this study, Mohammed also reported that the kappa statistic, which is a measure of the reproducibility between repeated assessments of the same categorical variable ranged between .79 and one for each item of the HHS and, according to the paper, indicated excellent reproducibility.

Note that both the Ragab and Mohammed studies did not include patient or physician evaluations of range of motion or deformity. These questions were eliminated from both the patient and the physician assessments.

Furthermore, McGrory and co-workers found that a brief follow-up phone call similar to the OOC follow-up procedure was effective in capturing missing data or clarifying multiple or contradictory responses from mailed patient self-assessment questionnaires.

In addition, the sponsor provided a

literature article by Barnes and co-workers which evaluated the reliability and validity of the OSHIP score as documented in their research paper. In Dr. Barnes' study, a group of 61 patients completed the OSHIP questionnaire. They were then sent a second copy to be completed two weeks later and returned by mail. The results of these two sets of surveys were compared to look for reproducibility. When comparing the responses the total interclass correlation coefficient was .93.

Dr. Barnes' study also included a separate group of 28 consecutive patients who were given both the patient administered OSHIP and a physiotherapist administered Harris Hip Score. The correlation between the patient's overall OSHIP score and Harris Hip Score was .91. The correlation between the individual corresponding domains ranged from .6 to .89, with the lowest correlation being for the patient's assessment of limp.

FDA requested additional correlations be provided that were not included in Dr. Barnes' study. In addition, FDA performed a linear regression

analysis to predict HHS score from OSHIP score for the 28 subjects. The linear regression analysis is summarized in your executive statistical summary, and the calculated R-Square is approximately .83.

A review of the raw data for the Barnes study also revealed the following. The average OSHIP score was lower than the Harris Hip Score, 62 and 67, respectively. Less subjects had an OSHIP score greater than 80 and more subjects had an OSHIP score less than 70 as compared to the Harris Hip Score.

There were 14 pairs of data where the OSHIP and HHS scores differed by more than five points. Of the 14 pairs, the HHS score was higher in 12 cases, while the OSHIP was higher in only two cases.

Additional information regarding the correlation and regression analysis and the limitations of the Barnes study will be summarized by our statistician, Dr. Chang Lao.

Like the Barnes study, Ragab also reported a relative lack of correlation between patient assessments of limp and the physician assessment of

limp, which he believed was due to a physician's tendency not to report limps that occurred only after long walks or during weather change, while patients were likely to report such limps.

However, unlike the Barnes study in which the OSHIP and HHS item regarding pain had a correlation of .83, Ragab found that when patients reported significant pain, they were often attributing the pain to their hips when the pain in most cases was not truly hip related. The authors reported that the physician was better able to distinguish true hip pain from pain coming from other sources, for example, secondary to trochanteric bursitis, lumbar spondylosis, and arthrosis of the contralateral hip.

An additional finding by McGrory and co-workers was that questions about whether patients could cut their toenails and put on shoes and socks correlated significantly with Harris Hip Score range of motion, with correlations of .57 and .53, respectively. The authors concluded that responses to these two questions could, therefore, be an estimate of the weighted Harris Hip Score range of motion.

Finally, Johnston and Smidt also reported that there is a distinct relationship between hip flexion and the question about shoe tying.

In the final comment of Dr. Barnes and co-workers' study, the authors stated that the Oswestry hip score is not intended to replace clinical evaluations at the critical phases following hip surgery, that is, at one year, five years, and ten years. However, it can be a useful tool along with X-rays to replace unnecessarily yearly follow-up following hip surgery.

The sponsor used the reference studies by Mohammed, McGrory and Barnes to justify the use of patient self-administered questionnaires to adequately report pain and function data.

Furthermore, the sponsor asserted that the close correlation of the overall OSHIP and Harris Hip Scores reported by Barnes and the tendency of the OSHIP scores to be somewhat lower relative to the Harris Hip Scores suggest that the OSHIP is a very close, although conservative, estimate of the Harris Hip Score.

Finally, for the purpose of the BHR study, an additional question about patient satisfaction was appended to the end of the OSHIP assessment questionnaire.

Now for the results. Procedures in the overall McMinn cohort were 70 percent men, 29 percent women, ages ranging from 13 to 86 years, with an average of 53 years. Ninety-one point nine percent of the procedures were [on patients] less than or equal to 65 years of age. The primary diagnosis was osteoarthritis in 75 percent, dysplasia in 15.8 percent, AVN in four percent, inflammatory arthritis in 2.4 percent, and other diagnostic indications in 2.7 percent.

All femoral head sizes were used in the overall McMinn cohort and almost all patients received either the standard cup or the dysplasia cup styles.

The follow-up rates for the X-ray/Oswestry combined cohort upon which most of the effectiveness analyses were performed are shown in this table. The follow-up rate at baseline was 80.6 percent and 90.8 percent at five years. There were 546 procedures

evaluated of the 601 expected at five years postop.

Now I'll present the safety data, and please note that one of the FDA questions will ask for your comments on whether or not the data contained in this PMA provide a reasonable assurance of safety.

There were a total of 27 revisions, which included ten revisions due to femoral neck fracture, six femoral head collapse, one dislocation, two AVN, and eight were revised due to infection.

Factors contributing to femoral neck fracture and head collapse included osteopenia, poor bone quality as evidenced by cysts in the femoral head and acetabulum, SLE, severe rheumatoid arthritis, infection leading to bone death, AVN, femoral cysts, and a malpositioned component.

There were a total of 2,912 adverse events in 1,669 of the 2,385 procedures for a rate of 70 percent. I believe that the sponsor has adequately summarized the adverse events in this study, except I want to comment on that it was reported that there were 589 procedures with a wound exudate for a rate of 25 percent. The sponsor stated that this was probably

due to a difference in reporting requirements.

There were 20 patient deaths in 26 procedures which the sponsor indicated were unrelated to the BHR device, and again, narratives were provided in the PMA for each of the patient deaths.

The sponsor also provided a metal ion literature analysis. The reference is in the PMA reported at serum and urinary metal ion concentration in patients with total hip replacements, with metallic components in general, and metal-metal articulating implants, in particular, increase in the post operative period. However, there does not appear to be conclusive evidence that elevated cobalt and chromium levels have detrimental effects in the total hip arthroplasty patients.

Now I'll present the effectiveness data. Please note that one of the FDA questions will ask for your comments on whether or not the data contained in this PMA provide a reasonable assurance of effectiveness.

The 1,626 procedures in the X-ray/Oswestry cohort contributed to the assessment of survivorship.

The estimated percent of procedures remaining free from revision at five years after the BHR procedure was 98.4 percent with a 95 percent confidence interval, having a lower bound of 97.3 percent and an upper bound of 99.5 percent.

The only marginally statistically significant difference in five-year survival probability was between patients with osteoarthritis at 98.8 percent and avascular necrosis at 92.1 percent as a primary diagnostic indication. Again, we evaluated survivorship for the X-ray/Oswestry combined cohort, as well as the McMinn patients, which included the additional McMinn patients.

And Dr. Chang Lao will present that information as well in FDA's analysis.

Regarding the radiographic data, three of the 108 procedures in the X-ray cohort for whom radiographs were available were radiographic failures at five years or 2.8 percent. One failure was due to a femoral radiolucency, one due to an acetabular radiolucency, and one due to both an acetabular radiolucency and a change in the acetabular

orientation of greater than five degrees.

For the pain and function data, the OSHIP score was used to evaluate the 1,111 unilateral procedures in the X-ray/Oswestry cohort. Mean OSHIP scores improve from 60 to 94.8 at five years. At first operative years two, three, four, and five the percentage of cases with good or excellent scores, that is, greater than 80 points, was 96.9 percent, 95.8 percent, 95.2 percent, and 92.8 percent, respectively.

For the patient satisfaction data at five years, 99.5 percent of the procedures were pleased or very pleased with the operation.

The sponsor submitted two literature controls. The D'Antonio reference included data on 514 Howmedica Osteonics, ABC, and Trident ceramic-on-ceramic total hip replacement procedures, and the Garino reference included data on 333 Wright medical ceramic transcend, ceramic-ceramic total hip replacement procedures.

In our review of these references, they appear to have significant differences as compared to

the data provided for the BHR device in this PMA, including different evaluations. The OSHIP score was used for the BHR and the HHS score was used from the literature. The length of follow-up, 18 to 36 months and two to four years for the controls with two to five years for the BHR study.

Mean baseline payment function scores. The mean baseline score was 60 in the OSHIP scoring system for the BHR study and 44 for the Harris Hip Score-Garino study, and it was not reported for the D'Antonio Study. There were also differences in the indications for use, including differences in the rate of dysplasia and AVN diagnostic indications.

Additional information regarding the literature controls was summarized by the sponsor and was contained in the panel packs.

Now, I'll discuss the applicability of the data collected outside the United States by a single investigator to the target U.S. population, practice of medicine, and U.S. orthopedic surgeon population.

Please note that one of the FDA questions will ask for your comments on this information.

Again, the clinical data were derived from a foreign clinical study conducted by a single investigator, Dr. McMinn, at the Birmingham Nuffield, but also six patients at the Little Aston Hospitals in the United Kingdom.

There was no racial or ethnic data, origin data, for the patients presented in the PMA. However, the sponsor provided the racial and ethnic distributions of the general U.S. and general U.K. populations and believes that they're similar.

There were noted difference in the higher percentage of people with African descent and other races in the general U.S. population as compared to the general U.K. population.

The sponsor also provided a comparison of the demographics and diagnostic indications for the BHR study and the literature reference by D'Antonio and co-workers, again, for the Howmedica Osteonics ceramic-ceramic device.

There are noted differences in the higher percentage of men, higher percentage of procedures with dysplasia, with 15.8 percent of the BHR study and

none reported for the ceramic-ceramic study; a higher percentage of inflammatory diagnostic indications; 2.4 percent for the BHR study and none for the ceramic-ceramic study; a lower percentage of procedures with AVN; 4.1 percent for the BHR study and 16 percent for the ceramic-ceramic study and post traumatic arthritis, none for the BHR study, and four percent for the ceramic-ceramic study.

The sponsor stated that the orthopedic practice of medicine utilized by Dr. McMinn is the same as the standard of orthopedic practice in the United States. The sponsor described Dr. McMinn's practice of medicine as follows. The operating room has laminar air flow and body exhaust suits. Dr. McMinn used a posterior surgical approach. Antibiotic prophylaxis was used interoperatively and for 24 hours postoperatively.

DVT prophylaxis using IV heparin interoperatively and compression stockings and low dose aspirin was used postoperatively for six weeks. Interoperative venting of the femoral shaft was used to prevent fat emboli.

Early ambulation, including full weight bearing with a walker on postop day one, hospital discharge at postop day six. After six weeks postop patients begin range of motion exercises. Postop activity for the first year is nonimpact or low impact and avoidance of high impact exercises.

Finally, the FDA advised the sponsor that the PMA may be subject to conditions of approval, including a post approval study to evaluate the long-term safety and effectiveness of the device. In response to FDA's advisory, the sponsor included a post approval study protocol, which included a nonrandomized, prospective, longitudinal, unblinded, multi-center trial to evaluate the long-term safety and effectiveness of the device. It included an enrollment of 150 patients at 15 sites. Inclusion and exclusion criteria were defined. Clinical and radiographic follow-up for five years; long-term follow-up assessment using a self-administered mail-in patient questionnaire for six to ten years, the questionnaire which would include three yes and no questions regarding patient satisfaction, whether they

have had a revision or a replacement, and expectation upon removal in the near future.

An analysis of the explanted device components and clinical and radiographic success and failure criteria were also defined. Please note that one of the FDA questions will ask for your comments on the proposed post approval study,

Now Dr. Chang Lao will briefly discuss the statistical information in the PMA.

DR. LAO: Good morning, panel members and audience here. I'm Chang Lao, Division of Biostatistics.

Today I'm going to present the first slide, summary of the patient accountability. The patient accountability, the previous speaker has already summarized in greater detail, and this slide is by a Oswestry study cohort and by unilateral and bilateral hip implant, total, 2,385 hips, and OSHIP score is only available for the X-ray and Oswestry cohort, not available in the McMinn cohort over a three-year time period.

My outline today basically constitutes

different parts. The first part is the basic statistical issues for the PMA. The second part is the summarized PMA statistical analysis, and the PMA statistical analysis included, first of all, survivorship, revision free analysis, and Oswestry/modified Harris Hip Score, and the correlation between HHS and OSHIP score, and it gives some summary.

Basic statistical issue for the BHR device is the first issue is the unique investigator, Dr. McMinn. No multi-center trial. So the question is how to generalize to all doctors, to all other centers, how to carry out training. That's a question based on the data from one doctor.

The second issue is no control group, BHR only. It is nonrandomized. It is combined retrospective and prospective registry data.

So the question is how to interpret results from this study to the general type of patient population, and sample size justification is neither prespecified hypothesis testing nor based on the confidence interval approach.

And the sponsor's post hoc justification based on constant revision rate over five years. Exponential distribution with some desired power, but probably if statistical justification cannot test that post hoc justification, usually in randomized trials samples should be prespecified beginning the study, in the product study design stage.

Continued basic statistical issues is on nonrandom patient selection. So the question is how generalized to a well defined group, subgroup of patients.

Again, the sponsor did a post hoc justification, complete demographics in terms of age, gender, diagnosis. Comparability of three study cohorts: X-ray, Oswestry and McMinn.

And another issue is unclear correlation between the OSHIP and the HHS scores, and no sample size justification between the subjects.

Nonrandom sample, 28 paired data, and there's no masking or order randomization. The masking and order randomization of the OSHIP/HHS data are very important because if I'm the patient, when I

rate my OSHIP score I should not be aware of the facility's rating. If I'm the physical therapist I should not be aware of the patient's self-evaluation. Otherwise I would be very easy to introduce a bias.

The second part is the PMA statistical analysis, a summary of the PMA data. The first study here is the survivorship, five year, and separated by three study cohorts. As you can see, all above 98 percent at five year for three study cohorts.

At the bottom of slide and the last line is the number of patients at the beginning of the study and the number of patients censored and the number of revisions by year. So at the end of five years, actually this is based on real data, not based on accountability from the number expect due or theoretically due or number expect.

If based on previous speakers, the number of patient accountability, the percentage of patient accountability, five years above 90 percent. But if based on real data, only about 21 percent complete five-year study of the total, 2,385 hips at the beginning of study.

So that's the difference between the real data and the hypothetical data.

Figure 2 is the number of revisions, 27 total revisions, two of the X-ray cohort is beyond five years. So about 25 revisions by cohort and the reason. And as you can see, total is 25 revisions out of 2,385 hips, about a 1.1 percent is relatively low. So sample size, you know, is justification to test the comparability of the Oswestry study cohort. It really has not enough power to detect the difference because not enough for the revisions, 25 or 27.

So one other question was how can you combine statistically justify the pooling of Oswestry study cohort, the X-ray, Oswestry, and the McMinn cohorts. There's basically three different tests. One is a log rank test, which the description is a log rank, is actually the compared of observed, expect number of revisions over time period, and you have optimum power if the three survival curves, a parallel issue added.

But you can see in the Figure 1 some crossover over time, not complete parallel.

Wilcoxon rank test assigns some more weight to earlier follow-up time when more patients at the risk of revision.

And lastly, the Cox proportion hazard model, which also requires parallelism of survival curves. The hazard rate, so-called hazard is a rate which is not a probability, not a proportion, which is average the number of events, revisions, per time interval. This says per year they have 25 total revisions divided by five. The average of five revisions per year, that's a hazard.

But the Cox proportional hazard, you need a parallelism of the survivor curves, and in the Figure 1 and the three curves are not completely parallel with the statistics over the years.

The results by statistical tests, the premise of doing this way statistically justifies pool of the study cohort into an overall conclusion, but at the bottom footnote there, as a statistical justification, there's only one of the requirements and which is not sufficient to justify pooling some other clinical technique required here to adjust the

pooling of data because those we study, the cohorts come from a different time period and maybe different data correction scheme. So you need some kind of other clinical input.

Anyway, the statistics here, the p-values now are not significant. Even if the parallelism assumption is violated, but because consistency of three different tests, so I would say that there's no statistical difference unless we start survivor curve of the three study cohorts.

This is the total number of complications by combined cohort and combined unilateral and bilateral hips over time, and you can see that year one the AVN, avascular necrosis has a very large number, don't automatically decrease. So the year one, 35, the total AVN 35 complications, and because not every complication will result in revision. So the total number of complications is much larger than total number of revisions, 25 revisions.

There's a secondary effective endpoint by year based on available data, combined cohort, and on unilateral hip only. We exclude the bilateral hip is

hard to evaluate OSHIP score.

As you can see at the baseline, 1,111 hips, total capital N, and the first column, second line, the small n, total number observed, 892. So the ratio of the small n and the capital N is 80 percent, and if you look at the common ratio, small n over capital N, you can see that year one and year two, you have about 25 percent of the hips not observed. They're missing here.

Then the missing data decreased at the year five, 91 percent of hips observed. What that means here is that the person missing one year, two, they reappear in the later year in the missing, and not a dropout. We call intermittent missing, intermittent missing, the missing to come back again.

So we should be very careful here to interpret the OSHIP score over time, and the assumption here is that you look at the name here, main OSHIP score 60 at baseline, dramatic improvement in the year five, right? About 95 percent. Standard deviation, standard error and a 95 confidence interval for the mean OSHIP score.

Ninety-five confidence interval very, very narrow because sample size is quite large. That's good, and the subject here is we assume OSHIP account between the patient date missed during the year, one year and other year to improve similarly as those patients who have complete data.

This assumption, unfortunately we call missing at random or missing completely at random, cannot test statistically, and so this assumption we should be very careful. Hopefully the missing data, year, one year, two, other year and another are not due to device or not due to complication but to something else. That's why some people have to be very careful.

So to put that table into this slide there, which is quite a medical improvement after year one, two to five year, but again, the missing data should be very careful.

The final issue is the correlation between the gold standard HHS/OSHIP score. The assessment in general probably is not randomized, only 28 paired data, no justification which is not based on

hypothesis test or confidence interval approach, and the OSHIP of a patient's self-evaluation, for instance, at the HHS by physio therapist from the same clinic.

So the question here is no masking, no order randomization. Timing I'm not sure what's the time distance between these different measurements. If you measure in the relevant time period, maybe more relevant than those times, they fly apart because OSHIP score can change.

So OSHIP has three major areas, HHS, four major areas. Those that score zero are the worst, 100 the best, and OSHIP pair function hip movement.

There's a summary of the correlation between HHS and OSHIP for each individual component, and the pain, function and the total, pain, function, the blue color here because they give them more weight and are more important than the other components.

The total score, as John Goode said, is about 0.91 correlation, 95 confidence interval. It's inside parenthesis there. So movement is a correlation moderately and otherwise the other I would

say correlation about .8 is a moderating pretty good.

But again, this correlation is subject to potential bias because of no masking and no randomization.

How can predict HHS from OSHIP? So we did the linear regression analysis, and you can see here the straight line is the linear equation, the intercept and slope at higher end range because the fitting is better, and better than the intermediate range.

An R-square, .83, 83 percent of variability of the data about the mean was explained by this equation. The square root of .83 is .9. That's the correlation for total score, as you can see from the previous table.

Now, we look at many ways to compare the OSHIP and HHS because the correlation is just one necessary condition, but not a sufficient condition for the prediction. So second analysis of complete mean score, the HHS mean score is 67 versus OSHIP of 62, a difference of five points, and that's the mean difference in 28 pairs. Still it is a 95 confidence

interval, 1.36 to 8.8, which doesn't include zero. So it would mean a HHS statistically higher than OSHIP mean score. P value is .01. That completes the mean difference on the total score.

Another comparison that can be done is to compare the binary account to be sure that the code of total OSHIP score based on excellent plus good. It means scores 80 and above. As you can see on this two-by-two table, a total agreement 80 or larger is eight plus 15, 23 pairs, of a total of 28 pairs, which is about 82 percent agreement.

And a probability, HHS larger or equal to 80 is 12 over 28 versus OSHIP, which is nine over 28. So HHS has higher proportion, 43 percent versus OSHIP, 32 percent, and you will compare, say this proportion differs a significant difference, which unfortunately cannot be tested because it has enough problems. In order to test the difference you will need a large number of prespecified study designs. You have enough number of discordant pairs, and here is only four and one, five discordant pairs, which are informative pairs and 23 agreement pairs that they don't give you

much information in testing the hypothesis of proportions.

So in summary. Basic statistical issues, the one investigator cannot statistically justify generalization of result to other physicians' centers. Post-hoc justification for sample size, patient selection, inclusion/exclusion criteria cannot statistically support post-hoc justification. No prespecified masking, no order randomization, no sample size determination for correlation analysis.

Some incomplete/missing OSHIP data. That 25 percent missing on the year one, two and nine percent year five. Assumption is similar in clinical results between complete and incomplete patients, as I say, missing at random, but again, we cannot statistically test that assumption.

And the life study is statistical conclusion based on available data. The mean age is about 53 years old. The range is 13 to 86, 92 percent above equal to 65. Twenty-seven revisions two beyond five years for the X-ray cohort.

Survival analysis of the Figure 1 of

everyone is above 98 percent, five year, and the mean OSHIP score at five-year, 94.8, and the percent excellent plus good, 92.8 percent at five-year, and the correlation, .91 total score.

So thank you. This is the end of my talk. Thank you.

MR. GOODE: This concludes FDA's presentation in the morning, and I'd like to acknowledge our team who worked on this, including Dr. Chang Lao, Patty Jahnes, Tracy Bourke, MaryAnn Wollerton, Mike Courtney, and from OSB Ronald Kaczmarek.

Thank you very much.

PANEL CHAIRPERSON NAIDU: Thank you.

I'd like to thank the FDA speakers for their presentations.

Does anybody on the panel have any questions for the FDA now? You may also ask the FDA questions this afternoon.

(No response.)

PANEL CHAIRPERSON NAIDU: Seeing none, let's break for lunch. We'll reconvene at 1:00 p.m.

Thank you.

(Whereupon, at 12:06 p.m., the meeting was recessed for lunch, to reconvene at 1:00 p.m., the same day.)

AFTERNOON SESSION

(1:03 p.m.)

PANEL CHAIRPERSON NAIDU: Dr. Jay Mabrey from Dallas will be opening this part of the meeting. He will help us focus our deliberations.

The panel will then deliberate on the information in the PMA and on the information the sponsor and FDA has presented this morning.

The panel can ask the sponsor and FDA questions at any time. After a general discussion, the panel will address the FDA questions. Then there will be a second open public hearing and FDA and sponsor summations.

Then the panel will conclude the deliberations by voting on the recommendation to the FDA concerning this PMA.

Dr. Mabrey.

DR. MABREY: Thank you.

Just to remind the panel, this is the disease that we're looking at today that's confined osteoarthritis of the femoral head. This particular image comes from Peter Bullough's excellent Atlas of

Orthopedic Pathology, who was also one of my professors at Cornell.

The purpose of my presentation this afternoon is to acquaint the panel with the clinical aspects of the device under consideration and to provide a perspective on its place in the orthopedic armamentarium.

Total hip versus hip replacement arthroplasty basically concentrates on two factors. Number one is that the femoral neck is replaced and that there is no femoral stem.

As you can see here, there's a substantial amount of bone left behind with the femoral replacement or the head replacement arthroplasty.

The proposed advantages, therefore, are that this bone is conserved, that it reproduces anatomic hip mechanics. There is greater stability as opposed to total hip replacement, and as you've heard earlier, easier revision to total hip arthroplasty.

The concepts of hip resurfacing include conservation of bone, sparing of femoral neck, optimizing stress transferred to the neck, and

enabling future revision. It also includes the placement of a large femoral head with a relatively thin acetabular component that usually relies on press fit fixation. It has a stable range of motion, and again, the purpose is to preserve normal hip biomechanics.

Contraindications to use, again, as you've heard earlier today, absolute contraindication should include the elderly with osteoporotic bone, metal hypersensitivity, and impaired renal function. Relative contraindications include inflammatory arthropathy, severe acetabular dysplasia, grossly abnormal geometry, and large areas of avascular necrosis.

The evolution of intelligent design of the hip resurfacing arthroplasty began with the Smith-Petersen Mold, which actually started out as a glass device implanted by Smith-Petersen in 1928, and I'll go into the details of that in a moment.

It then followed two paths of evolution. First was the path to hip resurfacing in which polyethylene was the bearing surface, and here is a

list of some of those individuals who participated in that path.

The second path was the path to metal-on-metal articulation. Again, a list of those participating in that path, with the result being a device similar to the one being presented today, resurfacing metal on metal.

The Smith-Petersen Mold, the first Vitallium prosthesis was implanted in 1938. As I said, he implanted a glass one in 1928. The particular case you see here was implanted in 1948, and that radiograph has a 46-year follow-up. The patient was functioning quite well and was symptom free, except for a slight limp on the left side.

The McKee-Farrar came along and now we're following the path of metal-on-metal technology and implanted several of these devices. This radiograph demonstrates 23-year follow-up after implantation and the authors suggest that there is approximately two millimeters of linear wear.

At this point in time, sphericity, clearance, and surface roughness were not necessarily

appreciated, and the success and longevity of individual components were probably more a result of chance than of design.

Muller introduced the metal-on-metal total hip arthroplasty in 1987. He implanted 18 surface replacements, 35 stem replacements, six of which were revised after functioning for up to 25 years.

Getting back to hip resurfacing, Paltrinieri and Trentani came out with this device in 1971, which is a thin walled, all polyethylene acetabular cup, and as we're all aware from the literature, it was prone to wear and developing significant osteolysis. The metal component was composed of stainless steel.

Wagner introduced his device in 1974. It was widely used in Europe, but again, the acetabular fitness here was only four millimeters.

The Tharies hip, the total hip articular replacement using internally centric shells was introduced in 1975 and had a variable thickness of 3.5 to 5.5 millimeters, but again, remember that the bearing surface is all polyethylene.

Dr. Harlan Amstutz designed the porous surface replacement, introduced it in 1983. The femoral head was a titanium alloy with mesh. The acetabulum was a titanium shell with a polyethylene liner, again, polyethylene in between.

Finally, in 1988, the Metasul metal-on-metal total hip arthroplasty was introduced. Larry Dorr was one of the early proponents of that device in this country and reported on 70 patients in the year 2000 using the cemented web or cup. He demonstrated a 94 percent survival rate at seven years and demonstrated no osteolysis with that particular device.

This is probably one of the earlier examples of metal-on-metal hip resurfacing introduced by Grigoris and Roberts, utilized hybrid fixation with an uncemented cup and a cemented head. They also introduced improved instrumentation for preparation of the femoral head and for sizing.

Right now this represents the world market for hip resurfacing with the Conserve Plus being introduced by Wright Medical Technologies in 1996, all

the way up to the Icon hip resurfacing introduced in 2004.

Four areas of interest that I'll discussion. Number one is one area that's associated specifically with hip resurfacing, and that is femoral neck fracture. Two and three are associated with metal-on-metal hip arthroplasty in general, and that includes acetabular fixation, as well as wear and metal ion concentration, and the fourth area of interest results from published U.S. studies of metal on metal hips resurfacing.

Femoral neck fractures and hip resurfacing arthroplasty are a result of a demanding surgical technique. They usually result from some type of femoral head defect or from an error in implantation.

This is the result of one of those that developed several months after implantation of the device.

The surgical technique itself is rather demanding. One must maintain careful angle of implantation of the femoral head. One must avoid notching the femoral neck and avoid impingement.

With regards to implanting the acetabulum, it's not quite analogous to total hip arthroplasty as the surgical exposure is somewhat more challenging when the femoral neck and portions of the head remain in place.

This demonstrates the close proximity of the femoral neck with the edge of the acetabular component, and if there are errors made an implantation impingement can occur.

In those cases of osteoarthritis resulting from femoral acetabular impingement, these areas may be compromised and may be prone to fracture.

Femoral head cyst formation is a problem in both osteoarthritis, as well as osteonecrosis. As this radiograph, again, from Dr. Bullough's Atlas demonstrates.

Femoral head defects are often encountered in implantation of femoral head resurfacing devices. This one demonstrates the appearance of the head after reaming for placement of the resurfacing implant. At that point the defects are filled with bone from the acetabular reamings.

One of our key concerns with regard to hip resurfacing arthroplasty studies which indicate that there is a difference in the stress distribution in hip resurfacing as opposed to total hips is you'll notice in the middle the total hip arthroplasty concentrates its load at the distal tip along the shaft and then distributes the rest of the load along the stem.

However, with his resurfacing arthroplasty as this study from Kuhl and Balle demonstrates from this year, the stresses tend to concentrate right at the femoral neck, and this has several implications.

Number one is stress shielding. On the left is a radiograph from Lilikakis' report in Orthopedic Clinics of North America this year, demonstrating a hip resurfacing one month out. If you look on the right, the same patient, the same resurfacing two years out, and there's a significant amount of thinning of that medial femoral cortex from stress shielding.

The other problem in this area is that of acetabular fixation. As with the other hip

resurfacing arthroplasties, the fixation of the acetabular cup is necessarily dependent upon press fit fixation. Most authors recommend under reaming by one millimeter and then protecting or at least controlling patient weight bearing for some time until there's osteointegration.

So the type of osteointegration and the type of technique used to insure stability is very important in a consideration of hips resurfacing arthroplasty as a device in the United States.

Another area of interest is that of metal on metal ion levels. As you can see from this study, from Clarke, et al., published in the Journal of Bone and Joint Surgery in 2003, they matched 22 patients with metal-on-metal hip resurfacing arthroplasty, with 22 patients with 28 millimeter metal-on-metal total hip arthroplasties, both with age, weight, and length after surgery. At a median of 16 months postop they found that the cobalt and chromium ion concentrations, the average concentrations for hip resurfacing arthroplasty were 38 and 53, respectively. The average ion concentrations for total hip were 22 and

18, respectively. Those differences were significant.

The black bars up there represent the maximum values seen, and if you can see it, down at the bottom is a small, green bar. That represents the upper limit of normal.

Manufacturing of these devices is key and is probably responsible for their renewed popularity and their increased longevity at this point. As I said before, earlier devices did not appreciate sphericity and clearance, and it was by chance that some of them lasted up to 25 years.

In this study from Rieker, et al., they looked at radioclearance versus run-in wear. The red star is outlining a 38 millimeter head with a 100 micron clearance between the ball and the cup, and you can see that the run-in wear is approximately ten.

When you increase that clearance now to almost 300, you basically increase the amount of run-in wear fivefold. This is appropriate for larger devices. Here's a 50 millimeter cup with approximately 140 microns of clearance, and again, if you increase the clearance in that same device, you

also increase the amount of run-in wear, thus increasing the amount of metal debris.

The U.S. experience with metal-on-metal hip replacement arthroplasty as a result of studies conducted by authors using IDE classified devices. This is probably the most well known by Harlan Amstutz that included 400 hips in 355 patients, IRB approved. Seventy-three percent of those patients were male. Note that the average age is only 48 years, but the range was from 15 to 70.

This was a hybrid metal-on-metal hip resurfacing arthroplasty. The Conserve Plus, there was no HA coating on the back side of the acetabulum, and this represents three and a half years' average follow-up.

Also, 43 patients in that series had dysplasia, representing 11 percent of the cases. Three-fourths of those were Crowe Type 1.

Survivorship. Dr. Amstutz is always highly critical and analytical of his own results, and he divided survivorship into two areas. The overall survivorship of this device at four years was 94.4

percent. However, if you divide the group up by what he termed surface arthroplasty risk index, those with a high risk index for failure only had an 88.8 percent survival rate, four years, versus those with a low surface arthroplasty risk index with a 97 percent survival rate.

The Harris Hip Score in spite of all of this was 93.5 on average at four years.

Just as a bit of explanation, the surface risk index consists of a maximum of six points. Two points are for femoral head cysts of greater than one centimeter. Two points are for weight less than 82 kilos. One point for previous surgery, and one point for a high activity level, and a risk index of greater than three was associated with a much higher revision rate.

So if one had one femoral head cyst and had one prior surgery or had a high activity level, one could be considered at high risk for further revision.

As I said, Dr. Amstutz is painfully honest about the results in his studies. This is a study

presented in the Journal of Bone and Joint Surgery in 2004 looking at 600 metal-on-metal hip replacement arthroplasties in which he reported five femoral neck fractures.

In this particular case, you see that there's a collapse of the cyst underneath the head of the prosthesis, and in this one, this was probably a procedural error. It's a little difficult to appreciate. However, the head was not fully seated because the pressurization of the cement did not allow the cap to come all the way down. This allowed a reamed area of the femoral neck to be exposed, and this patient suffered from femoral neck fracture.

In summary, metal-on-metal hip resurfacing arthroplasty is prone to high cobalt and chrome ion concentrations comparable to that of metal-on-metal total hip arthroplasty.

And, number two, femoral neck fracture is unique to this family of devices and deserves careful scrutiny with regards to appropriate patient selection and surgical technique.

We will note that devices similar to that

being considered for this PMA are currently being tested within the United States.

I'd also like to point out that femoral head resurfacing arthroplasty is not a standard procedure taught in U.S. orthopedic residency programs, and one has to look at whether widespread implementation of this technique would or would not reflect the results seen in countries where the procedure is more commonplace and may be part of their usual training program.

Thank you very much.

PANEL CHAIRPERSON NAIDU: Thank you, Dr. Mabrey.

Anybody on the panel have any questions for Dr. Mabrey?

(No response.)

PANEL CHAIRPERSON NAIDU: Seeing none, we should start a general panel discussion. We are asked to consider an unusual PMA based on a retrospective study designed by a single surgeon based on British data set. It is a challenging PMA. Nevertheless, I'd like to get input from all of the panel members, and

Dr. Mayor if you do not mind starting off commenting.

DR. MAYOR: I'd be happy to if I'm not held to too high a standard.

I'd like to go through the PMA as we received it, particularly with regard to the FDA review memo because there were several things in that review memo that might not have suggested that the writer had English as a second language, but there were some things that I was concerned about in terms of how that came out.

For instance, on page 4, the second paragraph describes a metal-on-metal resurfacing component with a high carbon content, but the high carbon content is assessed at 25 to 35 weight percent, which seems a little extreme.

I'm assuming that what that really meant to say was a .25 to .35 weight percent of carbon, and that a surface roughness was identified as greater than .05, which I think should have read less than. And if there's anything about that that I should be reinformed about, I'd be happy to hear about it.

The manufacturers may be better equipped

to address the question on page 6 which had to do with the screws locking into the cortical bone, but I understand that they also lock into the threaded lug during the final phases of their insertion. And when you're actually tightening the screw down, it becomes snugger and snugger because of the dimension between the two.

PANEL CHAIRPERSON NAIDU: Yes, would somebody from the sponsor like to respond to that?

MR. VELEZ-DURAN: Yes, I would like to. There's two questions. One is on page 4, a reference to high carbon content and surface roughness, and the other is about the screw locking and lug related to technique.

For the first question if I can get the attention of my colleague, I would like him to talk about the high carbon content and surface roughness.

MR. BAND: Tim Band, again, an employee of Smith & Nephew.

You're quite right. There were two typographical errors in the text. The carbon content is between .25 and .35 percent weight for carbon. It

was actually on my slide during the earlier presentation, and the surface references are maximum of .05 microns RA. So it should be less than rather than greater than.

PANEL CHAIRPERSON NAIDU: Does anybody else want to address the technique part?

MR. VELEZ-DURAN: I'm sorry. You may have to repeat the question on the technique.

MR. BAND: Tim Band again.

I think the question was about the use of the dysplasia screw in the lug on the acetabular dysplasia cup component, and the fact that the screw engages in the cortical bone, but then finally as it finally drives home in the acetabular component, the thread is also timed to be a full engagement.

The purpose for this is so that there's no leverage of the dysplasia cup component as the screw is driven further into the bone. It would have a potential levering of the cup over because it's as a timed position for the thread. It advances the cup in a perpendicular manner as opposed to any leverage.

Does that clarify?

DR. MAYOR: I think so.

Page 7 identifies the cobalt chrome beaded surface as a coating. I think if you are to be semantically rigorous, because it's a cast shape it's not a true coating.

MR. VELEZ-DURAN: The cast surface actually is a fully integral cast surface. So it is produced as a component part fully attached to the substrate and is, in fact, not a coating.

DR. MAYOR: It was on the basis of that perspective that abrasion testing for integrity was not done. Is that fair?

MR. BAND: In fact, we did do that, and that was submitted within the PMA package. So we have done testing of the porous surface.

DR. MAYOR: There was one specimen in the test protocol.

MR. GOODE: Dr. Mayor, if I could just clarify.

DR. MAYOR: I'm sorry. Yes, please.

MR. GOODE: My understanding is we did ask about abrasion. The sponsor, I believe, did provide a

justification for not doing it exactly like you just said, that the strength of that interface would be along the lines of the substrate of the metal because it is integrally cast. Therefore, that would not be required.

DR. MAYOR: Right.

MR. GOODE: So you're exactly correct.

DR. MAYOR: Yeah. Page 11-12, there was a discussion of the similar testing done on specimens of the implant system, and one component that was reported as having produced a wear rate which was considerably higher than the other four, but there was no convincing discussion of the assessment of why that might have occurred.

I'm concerned because we've had some retrievals at the laboratory up in Dartmouth which would suggest that that may be an occurrence with clinical significance.

MR. VELEZ-DURAN: Yes, I'd like to introduce Professor Unsworth, who is going to respond to that question.

DR. UNSWORTH: Thank you very much.

I'm Professor Tony Unsworth, Director of the Center for Biomedical Engineering at Stone University in the United Kingdom.

The university does receive research funding from Smith & Nephew, and Smith and Nephew covered my traveling expenses to come here today to be with you. Other than this, I have no financial interest in Smith & Nephew nor any of its products.

Yes, to try and answer the question you raised, sir, it is quite common in experiments of this sort. There are a number of reported incidents that the odd specimen, even though they're produced at the same specification as the rest do produce a high wear rate, and it has to do with the running in process. I'm afraid I don't know the exact mechanism, but it does happen from time to time, but normally they do restore themselves as this one did.

In fact, after about between three and five million cycles, the wear rate dropped very considerably, and that can be seen in terms of the surface asperities that were -- or the surface topography when we got to that stage.

I do have a slide if you're interested. I could show you the slide. I think it's about number six, something of that sort.

DR. MAYOR: Well, while I have you there, there is an issue related to surface topography that also came up on page 13 where the topography was measured, but could all be measured at the polar area because the lens of the instrument wouldn't fit into the cup.

How close to the edge of the lip could you get before you had to --

DR. UNSWORTH: You really only have the polar region in the cup because it's a noncontacting method of measuring the surface so that you don't damage the surface by putting a stylus across it, and so we have to get the lens into the cup, and it's fine on the head because, of course, it scans on the outside, but in the cup it was difficult to get it, other than in the polar region where the contact actually took place.

DR. MAYOR: Well, I'm wondering if you could get down perhaps to the Tropic of Cancer on

the --

DR. UNSWORTH: Sorry. If you could get down to?

DR. MAYOR: To the Tropic of Cancer on the cup if you couldn't get to the equator.

DR. UNSWORTH: I just don't know because I didn't do the experiment myself. It was one of my research associates. So I couldn't tell you how far down it could get. I apologize for that.

We've got the slide if you'd like to see what happened.

DR. MAYOR: If you've got them and could throw them up, that would be fine.

DR. UNSWORTH: Yes, thank you.

This was the first joint that didn't wear very rapidly, that wore at a nice, steady rate. So this is joint one, and when you friction test it this was the start of it.

And then after one million cycles, you can see that the surfaces were then becoming smoother. After two millions cycles, it looks like three million cycles is the next one please.

That's three million cycles, and most of them did that, except for that one that did not give the low pictures. So I'll just show you what happened with that one.

Could I have the next slide, please?

Again, started off very like the one. After four million cycles, it started to get smoother, but not as quickly as the others. Then again at three million cycles it was considerably rougher, but then at five million, by the time we got to five million cycles, we continued that on.

Yes, please, can we press it? There. It has become very like the original, which is now showing up there.

So it is smoothed down eventually, but it just took longer and not really in process.

PANEL CHAIRPERSON NAIDU: Thank you.

DR. UNSWORTH: Thank you, sir.

DR. MAYOR: And finally, I wonder. It has occurred to me to ask the clinicians in the group supporting this PMA. Why not do a resurfacing BHR on a 70 year old person, male or female, if bone stock

and habitus are favorable. Even with low activity on the list of intentions that the patient identifies is that because it lacks a track record or is that because there are some concerns about its use in those individuals who might be 70 or older, to expand on what it is that has inspired me to raise this issue.

My clinical experience in following patients who had their arthroplasties done at age 70 reminds me that I get very uncomfortable when I see them back at age 85 and they're starting to show lucent lines around their cement metal, and I'm beginning to think that I'm going to have to do a very difficult and stressful operation, which will be stressful not just for me but for them, and now that they're 15 years senior to the time at which I did their index total hip, and I guess the answer to that question might inform me a little bit better as to what would be the circumstances once this becomes widely available. Are there likely to be surgeons that would take the same approach in terms of for whom this approach might be beneficial since it's revisable, and revisability is not an insignificant

factor.

MR. VELEZ-DURAN: This is Marcos Velez with Smith & Nephew.

The first point of clarification I wanted to make is in the labeling that we discussed earlier. The activity level of the patient was a consideration in the selection of the patient. Age was one, but also the activity level of the patient was also as important.

But you asked for a clinician experience and comment. So I'd like Dr. Cecil Rorabeck to come up.

DR. RORABECK: Well, good afternoon, ladies and gentlemen. I'm Cecil Rorabeck, and I'm a conflict of interest as a consultant for Smith & Nephew and have been for many years. I consult with them on their hip systems.

I should also tell you that I've had conflicts with J&J and DePuy on the knee side and Zimmer as well on the knee side, but that is not relevant I don't think to what we're talking about today.

I'm a Professor and Orthopedic Surgeon at the University of Western Ontario in London, Canada, a member of the Hip Society in the United States, and the International Hip Society, American Academy of Orthopedic Surgery, and I restrict my practice to total hip and knee.

I have also been the past president of the Canadian Orthopedic Association, the Canadian Orthopedic Research Society, and currently am the Vice President of the College of Physicians and Surgeons of Canada.

So with all of that out of the way, let's try to deal with the question at hand, which is a good question, and I also am one who travels the world a lot like you all do, I'm sure, and looking at the probable indications of this procedure, it seems to me at least when you're starting out that you want to choose people with good bone stock and bone stock that's reproducibility good with time.

So what does that mean? Well, in my hands that means a male under the age of 65, probably, and it means a woman with a normal DEXA scan, probably

under the age of 60. Now, that's purely empiric. There's no scientific data to back it up, but if we accept the fact that with time people are more likely to become osteoporotic and if we accept the fact that the major shortcoming, if there is a major shortcoming, it's a potential for neck fracture. In my view at least when we are starting out, we should restrict the cases to patients under 65 or patients with normal bone stock, normal bone density.

Does that answer your question?

DR. MAYOR: In some respects, yes. I still wonder whether there's going to be a temptation because of revisability that this implant system may become increasingly attractive as a possible solution for a wider and wider array of patients.

DR. RORABECK: Well, I mean, you're quite right. It might happen that way, but you know something? We have such good, good things for patients, as you know, at age 75 to low demand. To me I don't think this is really what this implant is trying to address.

Now, if you have a 75 year old Swedish

farmer from Minnesota who has got fantastic bone stock and is very aggressive, perhaps there is a place for it, but I think it would have to be individualized in older patients.

DR. MAYOR: And the my final question is based at anyone who feels willing to approach it, and that is that we've dealt a lot with the issue of metal ion concentration in serum and urine, but I was concerned at the last academy meeting by my reading of a poster exhibit by Josh Jacobs who looked at a series of patients with bilateral, metal-on-metal total hips, and on the basis of the serum levels that he could identify in those patients during the subsequent measurement of their serum levels, he raised the concern that the levels for the bilaterals was more than twice the levels for the unilaterals, and that as he pursued the question in discussing the issues began to suggest to him that there might be a saturation problem that you could encounter in regard to clearance mechanisms that the body can use to deal with these ions, and that the load from a bilateral implant might actually exceed the capacity of those

clearance systems to respond to it.

And since there is some evidence provided just now by Dr. Mabrey that the serum levels may be higher for the resurfacing than they are for the smaller head metal-on-metals, does this raise a serious concern for those patients with bilateral implants and have we done any measurements which would either confirm or assuage the concerns that Dr. Jacobs is raising?

MR. VELEZ-DURAN: Marcos Velez from Smith & Nephew.

There's data on the PMA about the metal ion. They are all related to the BHR. However, to respond to your question to go into more details to response to your question, I would like to invite Mr. Joseph Daniel to the podium.

MR. DANIEL: Hi. I'm Joseph Daniel from Birmingham, England. I'm an orthopedic surgeon. I don't have any financial interest with Smith & Nephew, but my travel and stay here are being paid for by them.

Regarding the question, shall I take the

second part of the question first, which is that Birmingham hip resurfacings produce more hip resurfacings as a group, produce more metal ions than smaller diameter, 28 millimeter total hip replacement.

Now, if we look back at the fluid film lubrication theory, it would suggest that a larger diameter bearing has the potential to generate a full fluid film lubrication, and therefore is likely to wear less.

The article by Clarke which was shown earlier seemed to suggest that the resurfacings produce higher levels of metal ions.

Now, there was a confounding factor in that article and in that two types of resurfacings were combined, and the metallurgy and microstructure in that variable group is different in the two types of resurfacings, and in fact, that point has been highlighted by Dr. Josh Jacobs himself in his recent article in Journal of Arthroplasty in December 2004, that this confounding factor has been found.

And the work of Dr. Josh Jacobs himself in that article he presents, that he does not find a

difference between a 28 millimeter replacement and a larger diameter resurfacing.

Our own data also show that either in the 24-hour cobalt output or in the whole blood metal ion levels there is no significant difference either at the two-year period or at the five-year period between resurfacing and 28 millimeter metal-on-metal total hip replacements.

On the other issue about bilateral resurfacings and the question of renal threshold, the question of renal threshold can be looked at. When you look at the paper done some time ago in patients with renal failure and it's found that the metal ion levels in patients with renal failures tends to go up 100 times what it went up in regular people with no renal failure.

So the metal ion generation is not in terms of one or two times the metal levels in the blood, but much higher, and kidneys seem to have a large renal threshold to get rid of the excess metal ions.

We are, in fact, in the process of doing a

study of patients who had one resurfacing some time earlier and then come back for a quadrilateral hip resurfacing later on. We found in these patients that compared to the metal ion levels output in urine, daily output in urine, before the second of the quadrilateral hip operation, after the operation the levels go up more than three times in daily urine output.

The whole blood levels also go up, but they do not go up three times. They go up around twice the level before the second operation.

Now, there are differences between Josh Jacobs' technique and specimen and the technique and specimen that we have used. We have used whole blood rather than serum, and we have used high resolution and inducted a couple plasma mass spectrometry rather than rapid burning atomic absorption spectrometry.

Now, this point is also relevant. The reason why we chose whole blood rather than serum is because it has been shown in 1995 by Merritt and Brown that chromium especially tends to get sequestrated in blood cells, and so she has recommended that the